BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 9: e182514
doi: 10.3897/biss.9.182514

OPEN ACCESS

Conference Abstract

Mapping Nested Vegetation Plot Data to Darwin Core and the Humboldt Extension: Challenges and Open Questions

Salza Palpurina[‡,§], Kate Ingenloff[|]

‡ Global Biodiversity Information Facility, Europe and Central Asia Regional Support Team, Sofia, Bulgaria
§ National Museum of Natural History, Bulgarian Academy of Sciences, Sofia, Bulgaria
| Global Biodiversity Information Facility, Copenhagen, Denmark

Corresponding author: Salza Palpurina (spalpurina@gbif.org)

Received: 15 Dec 2025 | Published: 19 Dec 2025

Citation: Palpurina S, Ingenloff K (2025) Mapping Nested Vegetation Plot Data to Darwin Core and the Humboldt Extension: Challenges and Open Questions. Biodiversity Information Science and Standards 9: e182514. https://doi.org/10.3897/biss.9.182514

Abstract

Vegetation plots (relevés)—records of all plant species with their abundance in fixed-size plots in terrestrial vegetation—might be independent, arranged in nested designs, or established as permanent plots for resurvey. Given their complexity, representing vegetation plot data in global infrastructures such as the Global Biodiversity Information Facility (GBIF) requires thoughtful use of Darwin Core (DwC) and extensions (Darwin Core Task Group 2009). In DwC, a relevé corresponds to a single dwc:Event, hence, the recommended way to publish such data in GBIF is using the Event Core with at least two additional extensions: Occurrence and Relevé (GBIF 2018). The GBIF Relevé extension provides standard measurements across vegetation layers, with explicitly defined units and precision (e.g., “tree cover (%)”). Other plot-level measurements (e.g., soil pH) can be described via the DwC Measurement or Facts extension.

The recently ratified Humboldt extension (TDWG Humboldt Extension Task Group 2024) expands the DwC Event class to capture the sampling context of more complex plot designs by extending the descriptions of sampling events, making it especially relevant for nested and repeated sampling. However, recommendations are still evolving (Ingenloff et al. 2025) and it has not been extensively tested with vegetation plot

© Palpurina S, Ingenloff K. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

data (Suppl. material 2). Out of the 374 databases listed in the Global Index of Vegetation-Plot Databases (GIVD; accessed on 15.09.2025), only a few are currently published via GBIF (e.g., Hennekens 2018, Kuzemko et al. 2024, Swacha et al. 2025), and mostly as Occurrence class datasets.

As a case study, we present the dataset by Palpurina (2025) exported from Turboveg 2 (Hennekens and Schaminée 2001), the standard software for vegetation data management in Europe. For this dataset, we used an R script*[1] to directly access the database and export the data as Darwin Core Archive (DwC-A) using an Event Core with Occurrence, Relevé and Humboldt extensions. Regarding the Humboldt extension, field mapping required one-to-many and many-to-one relationships between Turboveg and DwC. Because some Turboveg 2 fields lacked sufficient detail, the script had to include hardcoded values and manual enrichment steps to populate missing DwC fields and ensure schema compliance. Our talk (Suppl. material 1) addresses challenges in applying the Humboldt extension to vegetation-plot data, focusing on common practices for capturing key data critical for vegetation data interoperability.

Keywords

biodiversity data standard, GBIF, relevé, sampling event

Presenting author

Salza Palpurina

Presented at

Living Data 2025

Acknowledgements

We thank the GBIF Secretariat for the financial support for publishing this work.

Conflicts of interest

The authors have declared that no competing interests exist.

References

- Darwin Core Task Group (2009) Darwin Core. Biodiversity Information Standards (TDWG). URL: http://www.tdwg.org/standards/450

- GBIF (2018) Macrophyte survey. In: GBIF (Ed.) Best Practices in Publishing Sampling-event data (version 2.2). GBIF Secretariat, Copenhagen. URL: https://ipt.gbif.org/manual/en/ipt/latest/best-practices-sampling-event-data#macrophyte-survey
- Hennekens S, Schaminée JJ (2001) TURBOVEG, a comprehensive data base management system for vegetation data. Journal of Vegetation Science 12 (4): 589-591. https://doi.org/10.2307/3237010
- Hennekens S (2018) Dutch Vegetation Database. Version 1.6. Wageningen Environmental Research. Sampling event dataset. URL: https://doi.org/10.15468/ksqxep
- Ingenloff K, Svenningsen C, Earl C, Shimabukuro PHF, Sica Y, Gan Y, Kachian ZR, Brenton P, Hochachka W, Wieczorek J, Stevenson R, Kazem A, Baskauf S, Zermoglio PF, Bloom D, Rodrigues A, Gamboa Martínez J, Schigel D (2025) Guide for publishing biological survey data to GBIF, version 1.0.2 (published 2025-09-24). GBIF Secretariat, Copenhagen. https://doi.org/10.35035/doc-ynvs-eh84
- Kuzemko A, Vynokurov D, Vasheniak I, Budzhak V, Chorney I, Chusova O, Dembicz I, Dengler J, Didukh Y, Dziuba T, Hájek M, Janišová M, Kolomiychuk V, Konaykova V, Moysiyenko I, Roleček J, Savchenko G, Shapoval V, Shyriaeva D, Škodová I, Tokariuk A (2024) Records of vascular plants, bryophytes and lichens from Ukrainian Grassland Database. Version 1.3. M.G. Kholodny Institute of Botany, National Academy of Sciences of Ukraine. Occurrence dataset. URL: https://doi.org/10.15468/w4yz9s
- Palpurina S (2025) Vegetation plots collected in dry grasslands throughout Bulgaria and Romanian Dobrudzha. Version 1.13. Masaryk University, Department of Botany and Zoology. Sampling event dataset. https://doi.org/10.15468/PKX4TG
- Swacha G, Kącki Z, Stefaniak J (2025) Polish Vegetation Database. Version 1.4. University of Wrocław. Occurrence dataset. URL: https://doi.org/10.15468/fzugec
- TDWG Humboldt Extension Task Group (2024) Humboldt Extension Vocabulary List of Terms. Biodiversity Information Standards (TDWG). URL: http://rs.tdwg.org/dwc/doc/eco/2024-03-26

Supplementary materials

Suppl. material 1: Presentation doi

Authors: Palpurina S, Ingenloff, K
Data type: PDF with notes
Brief description: Presentation delivered at the Living Data 2025 conference. This version was exported as a PDF including speaker notes, with minor modifications to some slides.
Download file (2.02 MB)

Suppl. material 2: Overview of vegetation-plot related datasets on GBIF doi

Authors: Salza Palpurina
Data type: dataset
Brief description: A summary table with datasets related to vegetation plots published on GBIF.org as accesssed via rgbif. Keywords, assembly and filtering are available upon request as an R script. Based on information in the table, as of 11.09.2025, GBIF hosted 524 vegetation plot-related datasets, of which 14 were metadata-only, 366 occurrence-only, and 144 were sampling-event datasets. Among the latter, only 22 employed the Relevé extension and only one also used the Humboldt extension.
Download file (1.17 MB)

Endnotes

*1 R script used to access directly the Turboveg 2 database and export selected vegetation plots as a DwC-A Sampling Event dataset with Occurrence, Relevé and Humboldt extensions is available at GitHub: https://github.com/spalp/GBIF-data-mobilization/blob/d39ec1b0e08af8cbc1ed5af029627f6465491c17/TURBOVEG-DwC_mapping.R